Regional Nonintrusive Load Monitoring for Low Voltage Substations and Distributed Energy Resources*

Shuangyuan Wang[a,b], Ran Li[b†], Adrian Evans[b], Furong Li[b]

[a]School of Mechanical Engineering,University of Shanghai for Science and Technology, Shagnhai, 200093, CN
[b]Department of Electronic & Electrical Engineering, University of Bath, Bath BA2 7AY, UK

Abstract

This paper presents a novel extension of the classic nonintrusive load monitoring (NILM) problem from household-appliance level to substation level. A new three-stage regional-NILM method is proposed to deduce the states of different types of loads in a region by disaggregating its substation demand. Three types of loads are considered in this study: i) traditional loads; ii) distributed generation such as photovoltaics (PVs); and iii) flexible loads like electric vehicles (EVs). The proposed method firstly forecasts the traditional load using the long-term historical data and employing spectral analysis to boost the signal-to-noise ratio. Secondly, the PV capacity is deduced by performing peak coincidence analysis between negative residuals and local solar irradiance data. Finally, a novel limited activation matching pursuit method is proposed to estimate the states of the EVs, including the total EV load and number of EVs. The method is assessed on real data collected from 800 substations, 10 PVs and 50 EVs in the UK. Results show the proposed method for estimating the number of EVs outperforms the approaches based on sparse coding, orthogonal matching pursuit and non-negative matching pursuit by 16.5%, 10.2% and 10.0%, respectively. The proposed Regional-NILM solution provides a cost-effective way for distribution network operators to understand the network's state. It can therefore significantly increase the network visibility without requiring expensive monitoring and avoiding data privacy issues. As such, it can improve the efficiency of demand side management, which is required to accommodate the future large number of distributed energy resources connections.

1. Introduction

To reduce its carbon emissions, the world is rapidly moving towards a low carbon future, through the roll-out of distributed energy resources (DERs) such as photovoltaics (PVs) and electric vehicles (EVs). In the UK, for example, new diesel and petrol cars will be banned from 2040. The connection of a large number of DERs, however, will have a significant impact on the power system. The distribution network will, for the first time, experience reversed power flow caused by distributed generation such as PV, which will present problems for existing network operations like voltage control and load forecasting. The additional peak demand from EVs will be far beyond the capacity of existing distribution network infrastructure, which was originally designed to supply kettles and lights. In the UK, the annual peak demand of EVs is estimated to surge from 5 GW to 30 GW, exceeding the capacity of Britain's distribution networks by 32% [1]. Such high demand will severely challenge the headroom and operation requirements of distribution networks. Many researches have investigated the impact of DERs [2] and proposed technical (e.g. energy storage) and commercial (e.g. smart pricing)

* The short version of the paper was presented at ICAE2018, Aug 22-25, Hong Kong. This paper is a substantial extension of the short version of the conference paper.

† Corresponding author. Tel.: +44 (0) 1225386183. *E-mail address:* r.li2@bath.ac.uk

innovations to mitigate the impacts via various solutions such as active distribution networks [3], virtual power plants [4], demand side management [5] and local energy markets [6].

A necessary condition for all these smart solutions is the visibility of loads, i.e. awareness of the number, type and state of the loads connected in a distribution network, so that the distribution network operators (DNOs) can accurately identify the causes of issues and thus control the appropriate DERs [7]. However, due to the lack of monitoring, DNOs actually have very limited visibility of the fast-growing DERs, including the capacity of PVs and states of EVs (e.g. number, type and location). It is impractical to install sensors on every DERs due to the associated costs and privacy issues. Therefore, it is critical to develop a cost-effective method to estimate the DERs connection states at a regional level without the need for individual monitoring, which is essentially a nonintrusive DERs monitoring approach for substations.

A cluster of similar research, termed non-intrusive load monitoring (NILM), has been widely studied to decompose the aggregated household energy into individual appliances' power consuming condition [8]. The concept was first introduced through the analysis of aggregated residential energy data, consisting of large power individual appliances with only ON-OFF states by matching the positive and negative changes of steady-states [9]. These techniques have been substantially developed over the last two decades by identifying new appliances such as finite state devices, continuously variable devices and permanent consumer devices [10]. New load signatures are proposed accordingly, including transient features such as edges and slopes [11], macroscopic signatures extracted from spectral analysis [12] and even the electric noise from a socket. Based on technical developments, the application of NILM has been extended from domestic houses to commercial buildings [13, 14], factories, data centres and EV charging stations [15]. Smart homes with energy storage, solar panel and EVs have also been studied [16, 17], which is of important practical interest in the energy management of smart grids. Beyond household devices, applications of NILM have been reported to remotely verify the 'health' of electromechanical loads and monitor power quality of electrical distribution systems [18]. The resolution of the data will affect the performance or accuracy of the NILM methods, especially when the sampling rate is low which is typically the case for smart metering devices being deployed worldwide where the sampling rate is in the order of minutes [19]. Therefore, it is very desirable to develop algorithms that can use smart meter sampling rate data. Such as real-time NILM [20], identifying appliance's anomalous behavior by NILM [21].

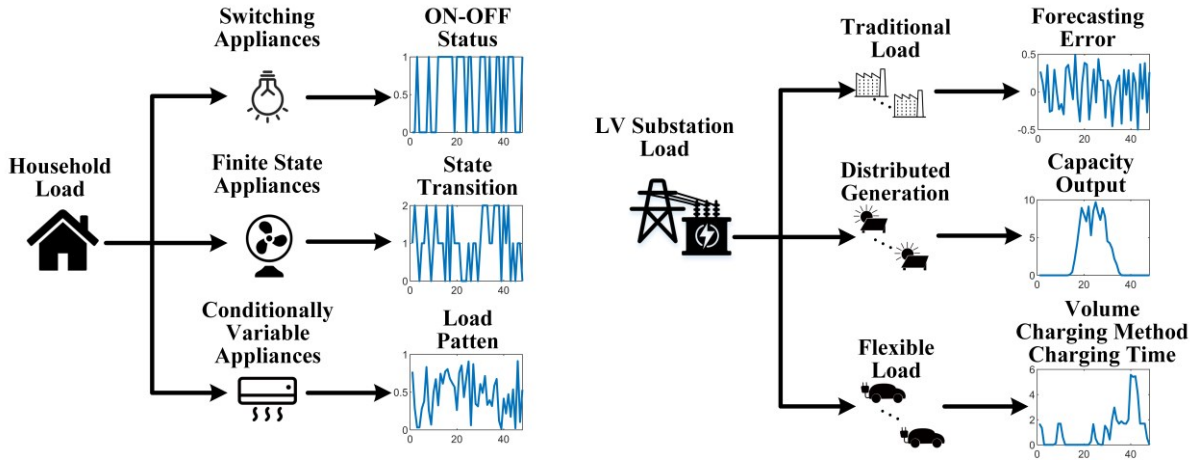

Fig. 1 (a) Traditional NILM at household level (b) Proposed regional-NILM at LV substation level

As shown in Fig. 1 (a), the majority of existing NILM techniques aims to deduce the switching time, state transition and energy pattern of the individual appliances within a household.

However, hardly any research has reported disaggregation of regional demand into different types of load for the purpose of increasing network visibility. Fig. 1 (b) illustrates this concept of decomposing a low voltage (11~0.415kV) substation's load into its components of traditional load, flexible load and distributed generation. The key challenges of applying existing NILM techniques to substation level are detailed below.

i) The traditional load can be estimated by classic load forecasting models to a reasonable accuracy based on historical data, customer and weather information. However, the residual may include a mixture of forecasting error, distributed generation output and flexible load. It is difficult to separate out these loads using traditional NILM methods when the signal-to-noise ratio is low.

ii) The outputs of distributed generators are equivalent to negative loads, which result in the total load having many possible disaggregation solutions, each consisting of different combinations of coincident loads and PVs. It is very difficult to apply traditional NILM methods without constraints if the solution is not unique.

iii) Unlike traditional loads, which are geographically connected at fixed points and temporally invariant in terms of quantity, the number of EVs connected to a substation is time-varying due to their mobility. Furthermore, their charging load patterns may vary significantly given the numerous different makes and charging modes. The unknown volume and charging patterns make it very complex to estimate EV states from the aggregated load.

To address these challenges, this paper proposes a novel Regional-NILM method which, for the first time, upscales existing NILM research from household-appliance level to substation level. The overall aim is to estimate the proportion and condition of the traditional load, distributed generation and flexible load for a given region based on aggregated substation data. The challenges are addressed through an advanced three-stage approach. The first stage is to forecast the traditional load based on the long-term historical data. In this step the forecasting residual is considered to consist of forecasting error, distributed generation and flexible load. As the main frequency of forecasting error is much higher than PVs and EVs, empirical mode decomposition (EMD) is adapted to filter out the forecasting error so as to increase the signal-to-noise ratio. The second stage is to deduce the PV capacity by performing association analysis between the negative component of the residual and local solar irradiation data. Given similar weather condition across the studied region, the PV pattern will be highly dependent on the meteorological data. The final stage consists of a novel limited activation matching pursuit (LAMP) [22] method to estimate the states of EVs, including the number of EVs connected to the substation and their charging load. Constraints are entered on the activation coefficients of the matching pursuit to avoid overmatching a particular type of EV. Unsupervised learning is adopted for the development of the basis based on typical EV charging profiles and the improvement of matching pursuit.

The remainder of this paper is organized as follows. Section II proposes the three-stage methodology of Regional-NILM. Section III explains each of the three stages on the technical details. Cases study are demonstrated in Section IV and conclusions are drawn in Section V.

2. Problem statement and the proposed regional non-intrusive load monitoring methodology

The overarching aim is to disaggregate the regional electricity demand into its traditional load, flexible load and distributed generation components. In this paper, EVs are considered as the flexible load and PVs as the distributed generation. The mathematical notation of all inputs are listed as follows:

I1) $\mathbf{L}_{sub} \in \square$, the aggregated load metered at the low voltage substation over the studied period, where n is the length of the data;

I2) $\mathbf{L}_{his} \in \square$, historical load metered at the substation, and N_d is the number of traditional customers, their load profiles, $\mathbf{P}_{trd} \in \square$;

I3) $\mathbf{H} \in \square$, the solar irradiance of the substation area;

I4) EV profiles extracted from sampled data.

The proposed method will deduce the following information as outputs:

O1) $\mathbf{L}_{trd} \in \square$, the amount of traditional load over studied region and period;

O2) C_{PV}, $\mathbf{G}_{PV} \in \square$, the PV capacity and total PV generation over the studied period;

O3) $\mathbf{L}_{EV} \in \square$, N_{ev}, $\mathbf{T}_{EV} \in \square$, the estimated total EV load, number of EVs and their charging profiles over the studied region and period.

This paper proposes a novel Regional-NILM methodology to decode the outputs through three stages as shown in the flowchart of Fig. 2. There are three stages in Regional-NILM, as described above. Each stage decomposes the aggregated substation load in to three components: traditional load, PV output and EV estimation. The decomposition of the aggregated substation load $\mathbf{L}_{sub}$, can therefore be expressed as

$$\mathbf{L}_{sub} = \mathbf{L}_{trd} + \mathbf{G}_{PV} + \mathbf{L}_{EV} . \quad (1)$$

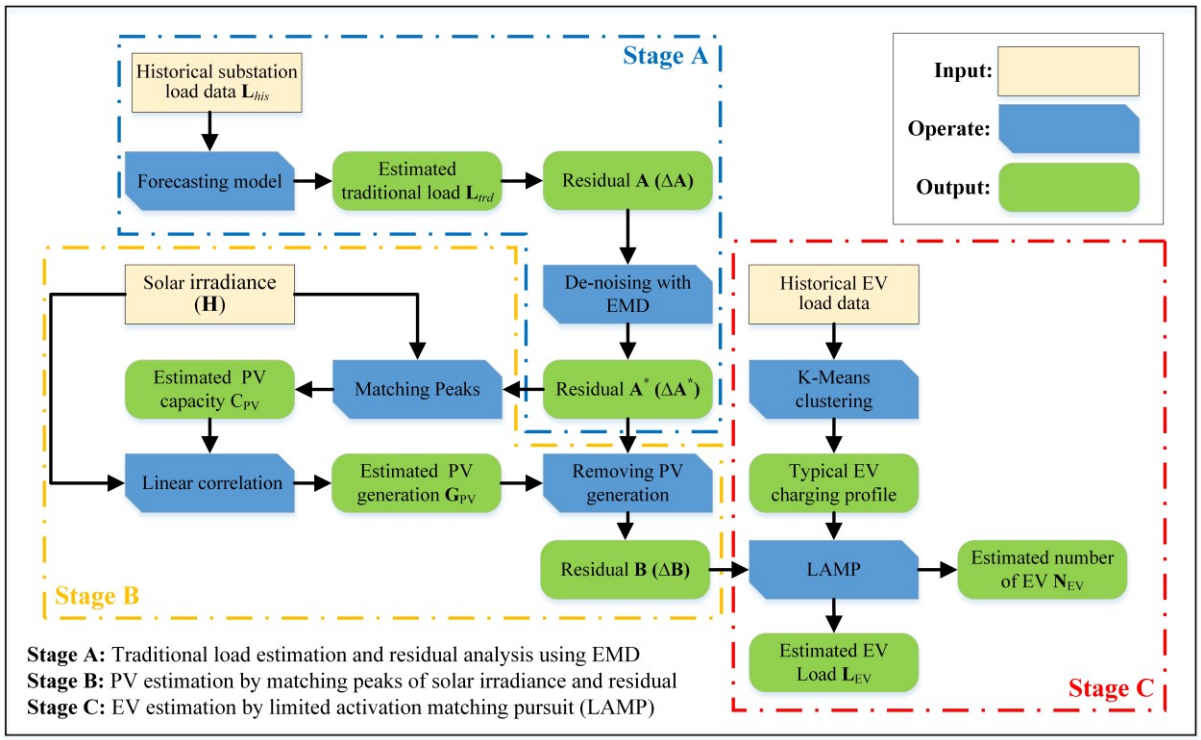

Fig. 2 The flow chart of the proposed regional-NILM method.

2.1 Stage A: Traditional load estimation and residual analysis using empirical mode decomposition

Stage A estimates the traditional load component using forecasting models. Here, two types of forecasting models are used to estimate the traditional load. For regions where traditional load has dominated in the past the substation's historical load data can be considered as DERs free. For regions where a good number of EVs and PVs have already been connected, load profiles and geo-demographical information can be used to eliminate the effects from existing PVs and EVs. For example, in the UK, DNOs uses 8 typical load profiles developed by Elexon [23] to estimate traditional customers. In this study, time-series forecasting models can be used to estimate the traditional load based on historical load data as very low penetration of DERs can be seen in the trial areas [24]. Existing forecasting models such as ARIMA, exponential smoothing and TBATS [25] can accurately predict the traditional load. In this study, a range of forecasting models have been tested on our data. While the deep-RNN model [26] performed best, its training stage is time-consuming. Therefore, the TBATS model is chosen as it gives very similar performance with a significant reduction in computation time.

It is reasonable to assume an unbiased prediction and thus the prediction residual follows the normal distribution. Given the estimated traditional load is $\mathbf{L}_{est}$, and the actual traditional load is $\mathbf{L}_{trd}$, the *estimating error*, $\mathbf{Err}_{est}$ can be expressed as

$$\mathbf{Err}_{est} = \mathbf{L}_{trd} - \mathbf{L}_{est} . \tag{2}$$

Subtracting the estimating traditional load $\mathbf{L}_{trd}$ from the total substation load $\mathbf{L}_{sub}$ produces *Residual* **A**, which consists of the EV load $\mathbf{L}_{EV}$, PV generation $\mathbf{G}_{PV}$, and the estimating error of traditional load, $\mathbf{Err}_{est}$. The *Residual* **A** ($\Delta\mathbf{A}$) can be expressed as

$$\begin{aligned} \Delta\mathbf{A} &= \mathbf{L}_{sub} - \mathbf{L}_{est} \\ &= (\mathbf{L}_{trd} + \mathbf{G}_{PV} + \mathbf{L}_{EV}) - \mathbf{L}_{est} . \\ &= \mathbf{Err}_{est} + \mathbf{G}_{PV} + \mathbf{L}_{EV} \end{aligned} \tag{3}$$

Even if the forecasting is performed with state-of-the-art models, an estimating error >3% is expected. If the estimating error is considered as noise, the resulting signal-to-noise ratio (SNR) of *Residual* **A** is low, which makes it difficult to apply traditional NILM methods to decompose it into the EV and PV components. As the forecasting error is nearly white noise while the EV and PV components show strong periodicity (daily for PV and 1-7 days for EVs), spectral analysis is adopted to reduce the high-frequency components of *Residual **A**,* thus increasing its SNR. In this study, EMD [27, 28] is used to denoise the signal and hence reduce the *forecasting error*, $\mathbf{Err}_{est.}$ The EMD method can decompose non-stationary data set into a finite and often small number of components. These components are described as intrinsic mode functions (IMF) and form a complete and nearly orthogonal basis for the original signal [28]. There are two important criteria that must be satisfied if IMFs are to be successfully employed: i) the number of extrema and the number of zero crossings must either be equal or differ at most by one; ii) the mean value of the upper and lower envelopes should be zero [29].

IMFs are generated from the original signal $\Delta\mathbf{A}$ using the Algorithm 1 process.

Algorithm 1. The pseudo code of EMD decomposition.

1. Input: $\boldsymbol{r}_0 = \Delta\mathbf{A}$, $i = 1$;
2. Extract the i-th *IMF*:
   (a) initial: $\boldsymbol{h}_0 = \boldsymbol{r}_{i-1}, j = 1$;
   (b) find all local maxima and minima of $\boldsymbol{h}_{j\text{-}1}$;

(c) use two cubic spline lines to fit the maxima and minima as the upper and lower envelopes;

(d) calculate the mean value $\boldsymbol{m}_{j\text{-}1}$ of the upper and lower envelopes;

(e) $\boldsymbol{h}_j = \boldsymbol{h}_{j\text{-}1} - \boldsymbol{m}_{j\text{-}1}$;

(f) if $\boldsymbol{h}_j$ is the IMF function, $IMF_i = \boldsymbol{h}_j$; else, $j = j+1$, repeat from (b).

3. $\boldsymbol{r}_i = \boldsymbol{r}_{i-1} - IMF_i$;
4. if the number of extremum points for $\boldsymbol{r}_i$ are more than 2, return to step 2, and $i = i + 1$; else, $\boldsymbol{r}_i$ is the residual component.

A total of n IMFs are obtained in this process with a final residual $\boldsymbol{r}_n$. The *residual A* ($\Delta\mathbf{A}$) can thus be rewritten as

$$\Delta\mathbf{A} = \sum_{i=1}^{n} IMF_i + \boldsymbol{r}_n . \tag{4}$$

The frequencies of IMFs decrease from IMF_1 to IMF_n. Since the first *u IMFs* usually carry the most oscillating (high-frequency) components, the *forecasting error*, $\mathbf{Err}_{est}$, is mainly distributed in this part. The specific value of u is determined by the frequency components of the IMFs. In general, the high-frequency components are mainly concentrated in the first IMF. Hence, in order to reduce the loss of information in the reconstructed signal, the value of u=1 is selected. Reconstructing the remaining low-frequency IMFs and residual $\boldsymbol{r}_n$ obtains the denoised *Residual* $\mathbf{A}^*$, $\Delta\mathbf{A}^*$, which is given by

$$\Delta\mathbf{A}^* = \sum_{i=u+1}^{n} IMF_i + \boldsymbol{r}_n . \tag{5}$$

As $\Delta\mathbf{A}^*$ is considered to only include the PV and EV components, it provides an estimate of the DERs load.

2.2 Stage B: PV estimation by matching the solar irradiance and residual peaks

A widely used formula for PV output estimation [30] is

$$\mathbf{E} = A \cdot R \cdot \mathbf{H} \cdot P , \tag{6}$$

where $\mathbf{E}$ is the PV generation in terms of energy (kWh), A is the total solar panel area (m^2), R is solar panel yield efficiency (%), $\mathbf{H}$ is the half hour average solar irradiance on tilted panels and P is the performance ratio, a coefficient for losses.

It is assumed in a short period of time (10 days in this study), the total PV capacity (i.e. installed area) of the solar panel is a constant, and the other coefficients are relatively stable. As the amount of PV generation is proportional to the intensity of solar irradiance, the PV output can be rewritten as

$$\mathbf{G}_{PV} = K \cdot \mathbf{H} , \tag{7}$$

where K is the comprehensive conversion coefficient between the solar irradiance and PV generation and $\mathbf{G}_{PV}$ is the average output power of the PV. In this paper, peak coincidence analysis is proposed to estimate the value of K for the studied area and period. The idea is to search for a period when a local minimum of the negative part of *Residual* $\mathbf{A}^*$ ($\Delta\mathbf{A}^*_-$) coincides with the local maximum of solar irradiance $\mathbf{H}$. The estimated value of the coefficient K can be described as

$$\hat{K} = \frac{1}{Z_{\alpha=1}} \sum_{i=1}^{p} \sum_{j=1}^{q} \alpha (\Delta \mathbf{A}_{-}^{*})_{i}^{idxi} / (\mathbf{H})_{j}^{idxj}, \ \alpha = \begin{cases} 1 & |idxi - idxj| \leq \grave{o} \\ 0 & others \end{cases}, \quad (8)$$

where *idxi* is the index of the *i-th* local minimum value of $\Delta \mathbf{A}_{-}^{*}$ and *idxj* is the *j-th* index of the local maximum solar irradiance **H**. It is noted that $idxi, idxj \in [1, n]$, where *n* is the length of the signal. The coefficients *p* and *q* are the numbers of peaks for $\Delta \mathbf{A}_{-}^{*}$ and **H**, respectively. $(\Delta \mathbf{A}_{-}^{*})_{i}^{idxi}$ is the value of the *i-th* peak of $\Delta \mathbf{A}_{-}^{*}$, and $(\mathbf{H})_{j}^{idxj}$ is the value of the *j-th* peak of **H**. α is the matching indicator, which equals 1 when the two peaks are matched and equals 0 otherwise. When the difference between the index values of the two peaks is smaller than a non-negative integer number ò, $|idxi - idxj| \leq \grave{o}$, two peaks are considered to be matched (if $\grave{o} = 0$, the two peaks are exactly matched, otherwise, the two peaks differ in time by ò times the sampling period). $Z_{\alpha=1}$ is the number of matched peaks.

The input to *Stage B* is the signal after the low-pass filter, *Residual* **A***, only consisting a mixture of EV load and PV generation. Since the capacity of PVs and the number of EVs are unknown, there are still infinite possible combinations of EVs and PVs to form the target load. For example, a load of 2 kW might come from 2 kW of EVs or the sum of 10 kW of EVs and 8 kW of PV. It is impossible to determine the number solely based on the load data. Taking advantage of the positive correlation between PV generation and solar irradiance, *Stage B* will first estimate the PV capacity. It is assumed that if the local maximum of solar irradiance coincidences with the local negative minimum of the *Residual* **A***, the coincident period is dominated by the PV generation. The total capacity can be therefore estimated according to

$$C_{PV} = \beta \cdot \hat{K}, \quad (9)$$

where C_{PV} is the total capacity of PV in the studied area and β is a coefficient representing the relationship between the PV output energy and the solar irradiance, which is estimated from the historical PV generation data.

The energy generated by PV can then be estimated from the solar irradiance and comprehensive conversion coefficient using

$$\hat{\mathbf{G}}_{PV} = \hat{K} \cdot \mathbf{H}. \quad (10)$$

The *Residual* **B**, $\Delta \mathbf{B}$, is obtained by subtracting the PV generation from the *Residual* $\mathbf{A}^{*}$ ($\Delta \mathbf{A}^{*}$) such that

$$\Delta \mathbf{B} = \Delta \mathbf{A}^{*} - \hat{\mathbf{G}}_{PV}. \quad (11)$$

The *Residual* **B** is the initial estimate of the EV load $\hat{\mathbf{L}}_{EV}$, and will be used as the input signal for *Stage C*.

2.3 Stage C: EV estimation by limited activation marching pursuit

In *Stage C*, the first challenge of EV load estimation comes from the uncertainty of charging modes. Unlike other appliances which have a unique profile, EVs have multiple charging modes such as slow charging, fast charging and extra fast charging [31]. We therefore need to identify the number of EVs and the charging time from the estimated load. Matching pursuit (MP) [32] is a greedy method for decomposing signals based on a redundant dictionary. The redundant dictionary is denoted by the matrix $\mathbf{D} \in \square$, in which every column of **D** ($\mathbf{d}_k \in \square$, $k = 1, \cdots$) is a prototype signal called atom or basis, *T* is the length of the atom

and p is the number of atoms. A given signal $\mathbf{x} \in \square$ can be represented as a linear combination of these atoms

$$\mathbf{x} = \mathbf{DA} + \boldsymbol{\gamma}, \tag{12}$$

where $\mathbf{A} \in \square$ is a sparse matrix indicating the scalar weighting factors for the atoms, which is also known as the activation coefficients and $\boldsymbol{\gamma} \in \square$ is the residual signal. Sparse **A** means that not every atom in **D** will be activated. Normally, MP selects the best single atom that gives the maximum reduction in the approximation error. The minimum stable error will be achieved through many iterations of this step.

In this study, EVs with different charging modes are treated as different appliances. The K-Means cluster method [33] is used to extract typical load profiles for different EV charging modes. Atoms are selected by removing those charging profile data with high correlations from the dictionary. The correlation coefficient is defined by

$$\rho(\mathbf{d}_i, \mathbf{d}_j) = \mathrm{E}[(\mathbf{d}_i - u_i)(\mathbf{d}_j - u_j)]/(\sigma_i \sigma_j), \tag{13}$$

where $\mathbf{d}_i$ and $\mathbf{d}_j$ are two atoms with expected values u_i, and u_j and standard deviations σ_i and σ_j and $\mathrm{E}[\bullet]$ is the expectation operator. Any charging profile will be removed if its correlation with an existing atom is higher than the defined threshold τ. For some EVs, no charging activities are observed throughout a day. Initially, a threshold value e is defined to eliminate those atoms with low-value noise over the whole periods. A dictionary $\mathbf{D} \in \square$ is constructed with the EVs load profiles

$$\mathbf{D} = [\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_p], \tag{14}$$

where $\mathbf{d}_i \in \square$ is the charging normalized profile of the *ith* EV. For a 24 hours length charging profile with half-hourly sampling cycles, the length of the atom is n = 48. The normalized energy is constrained by $\|\mathbf{d}_i\|^2 = 1$ and $i = \{1, 2, \cdots$, where p is the number of atoms in the dictionary. To reduce the end effects, the substation load profile $\mathbf{x} \in \square$ is then expanded to a new vector $\mathbf{y} \in \square$ with interpolated zero vectors, $\mathbf{y} = [\mathbf{0}^n, \mathbf{x}^m, \mathbf{0}^n]$, where $\mathbf{0}^n = [0, 0, \ldots, 0]$ is a zero vector with n elements, and m is the length of **x**.

The second challenge comes from the uncertainty of charging locations. EVs are a mobile load and the number of EVs connected to a substation is time-varying. *Stage C* proposes a limited activation matching pursuit (LAMP) [22] to estimate the number of EVs by decomposing *Residual B* into some combinations of typical EVs charging profiles. As the start point of EV charging may not match the start point of *Residual B*, the shift-invariant [34] method is adopted to improve the conventional matching pursuit, which reconstructs the signal using all the atoms in all possible shifts. Then, a limited activating constraint is applied to the convolution of the signal and the dictionary to account for practical constraints including: i) maximum charging power of EVs; ii) the maximum number of EVs in the area; and iii) non-negative activation coefficients as currently most EVs cannot discharge back to the grid.

The shift convolution of dictionary **D** and signal **y** is the target matrix $\mathbf{T} \in \square$, which will be used to perform the matching by

$$\mathbf{T} = conv(\mathbf{D}, \mathbf{y}), \tag{15}$$

where *conv* is the convolution operator. **T** is a $p \times (2n + m)$ matrix,

$$\mathbf{T}=\begin{pmatrix} T_{1,1} & \cdots & T_{1,(2n+m)} \\ \vdots & \ddots & \vdots \\ T_{p,1} & \cdots & T_{p,(2n+m)} \end{pmatrix}. \tag{16}$$

Hence, the activation coefficients can be expressed as the matrix

$$\mathbf{A}=\begin{pmatrix} A_{1,1} & \cdots & A_{1,(2n+m)} \\ \vdots & \ddots & \vdots \\ A_{p,1} & \cdots & A_{p,(2n+m)} \end{pmatrix}. \tag{17}$$

At the beginning of the matching, $\mathbf{A}$ is initialized as zero matrix, $\mathbf{A}^{(0)}=\mathbf{0}$, giving $\mathbf{T}^{(0)} = \text{conv}(\mathbf{D}, \mathbf{y}^{(0)})$, where $\mathbf{y}^{(0)}$ is the original signal of $\mathbf{y}$. The maximum value of $\mathbf{T}^{(0)}$ is $T_{max}^{(0)} = T_{i,j}^{(0)}$, where, i indicates the atom index, and j indicates the time index. The value of the activation coefficient in the first iteration is

$$a^{(1)}=\begin{cases} \ell & T_{max}^{(0)} > \ell \\ T_{i,j}, & T_{max}^{(0)} \le \ell \end{cases}, \tag{18}$$

where ℓ is the maximum allowable value under the limited activation mechanism.

To update the activating coefficients matrix $\mathbf{A}^{(1)}$, we only need to update one element in $\mathbf{A}^{(0)}$ giving

$$\mathbf{A}^{(1)}=\begin{pmatrix} A_{1,1}^{(0)} & \cdots & A_{1,(2n+m)}^{(0)} \\ \vdots & & \vdots \\ A_{p,1}^{(0)} & \cdots & A_{p,(2n+m)}^{(0)} \end{pmatrix}. \tag{19}$$

$\mathbf{y}^{(1)}$ is defined as

$$\mathbf{y}^{(1)}=\mathbf{y}^{(0)}-a^{(1)}\square \quad . \tag{20}$$

After one iteration, the convolution matrix $\mathbf{T}^{(1)}$ is given by

$$\begin{aligned} \mathbf{T}^{(1)} &= conv\left(\mathbf{D}, \mathbf{y}^{(1)}\right) \\ &= conv(\mathbf{D}, \mathbf{y}^{(0)} - a^{(1)}\square \quad , \\ &= \mathbf{T}^{(0)} - a^{(1)}\square \quad \mathbf{d}_i\big) \end{aligned} \tag{21}$$

Here, using the definition

$$\mathbf{G}(i) = conv(\mathbf{D}, \mathbf{d}_i), \tag{22}$$

gives

$$\mathbf{G} = conv(\mathbf{D}, \mathbf{D}), \tag{23}$$

where $\mathbf{G}$ is the convolution between the atoms in the dictionary $\mathbf{D}$, i is the index of the activating atom in the dictionary which has the maximum activating value as described above. Likewise, after k+1 iterations, the coefficients are updated by (24) - (28) as follows:

$$a^{(k+1)}=\begin{cases} \ell & T_{max}^{(k)} > \ell \\ T_{i,j}^{(k)}, & T_{max}^{(k)} \le \ell \end{cases}, \tag{24}$$

$$\mathbf{A}^{(k+1)} = \begin{pmatrix} A_{1,1}^{(k)} & \cdots & A_{1,(2n+m)}^{(k)} \\ \vdots & & \vdots \\ A_{p,1}^{(k)} & \cdots & A_{p,(2n+m)}^{(k)} \end{pmatrix}, \tag{25}$$

$$\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} - a^{(k+1)} \mathbf{d}_i, \tag{26}$$

$$\mathbf{T}^{(k+1)} = \mathbf{T}^{(k)} - a^{(k+1)} \mathbf{G}(i), \tag{27}$$

$$T_{max}^{(k+1)} = T_{i,j}^{(k+1)}. \tag{28}$$

As the activation only affects a particular range of $\mathbf{T}$, we only need to update the activated range $\mathbf{r} \in [\,]$ in $\mathbf{T}$. If n is even, $\mathbf{r}$ could be derived by $\mathbf{r} = [-n/2+j, -n/2+j+1, \ldots, n/2+j]$, else $\mathbf{r} = [-(n+1)/2+j, -(n+1)/2+j+1, \ldots, (n+1)/2+j]$, where j is the time index of the activation. Therefore (27) can be updated to give

$$\mathbf{T}(:,\mathbf{r})^{(k+1)} = \mathbf{T}(:,\mathbf{r})^{(k)} - a^{(k+1)} \mathbf{G}(i). \tag{29}$$

The iterations terminate when either the maximum iteration number N is reached, or the residual signal is smaller than a threshold value.

Fig. 3 shows the detailed scheme of the limited activating matching pursuit and its four steps are described below:

Step 1: the target matrix $\mathbf{T}$ is found by the convolution of the dictionary with the expanded signal, and the dictionary convolution matrix $\mathbf{D}$.

Step 2: the activation coefficients matrix $\mathbf{A}$ is updated by comparing the maximum point in the target matrix $\mathbf{T}$ and the limited activation threshold ℓ using (24) and (25).

Step 3: use (29) to update the activation range $\mathbf{r}$, which is calculated in the target matrix $\mathbf{T}$ as is best matching atom index i in the dictionary.

Step 4: check as to whether the termination condition is reached. If not, repeat steps 2 and 3.

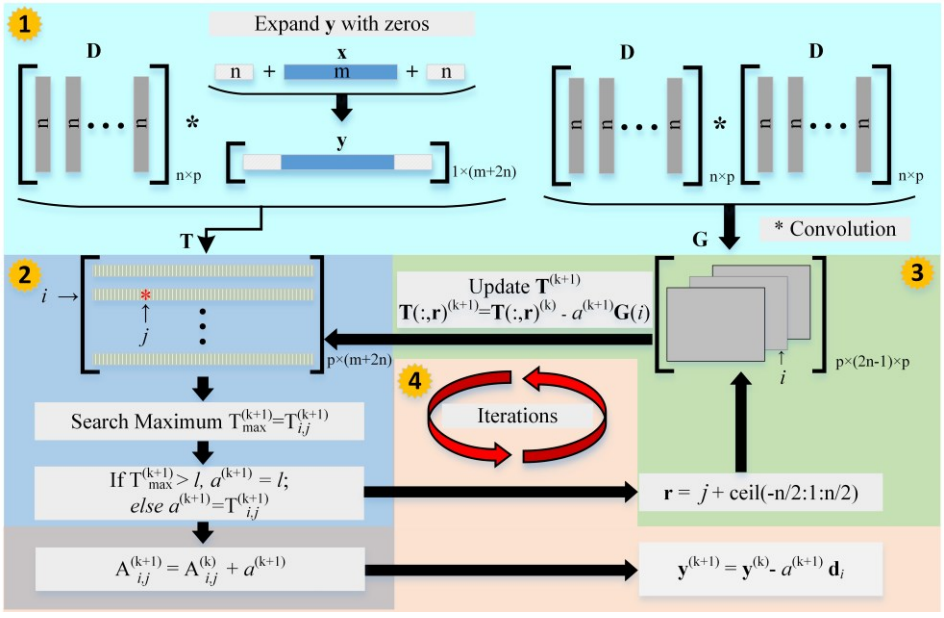

Fig. 3 Scheme of limited activating matching pursuit.

It is noted that EVs' charging profiles are discrete. There are some zero periods in the atoms indicating that the EV is not connected at the time. Therefore, the number of activating EVs should be calculated for each time period. The status of charging in an atom can be written as

$$\mathbf{z}_i(k)=\begin{cases}1, & \mathbf{d}_i(k)>\varepsilon\\ 0, & \mathbf{d}_i(k)\le\varepsilon\end{cases}. \tag{30}$$

Here, ε is a small value relative to the minimum charging power of EV. k is the index in the atom, k=1, 2, …, n. So, the number of EVs connected to the charging station N(t) at a particular time t could be expressed as

$$N(t)=\sum_{i\in S,\,t-n+1\le s_i\le t}\mathbf{z}_i(t\text{-}s_i\text{+}1), \tag{31}$$

where, s_i is the active time of the *ith* atom, S is the sets of active atoms index, and n is the length of the atom.

3. Implementation of regional-nonintrusive load monitoring

In this section, the proposed method is implemented on real data collected from the UK. For traditional loads, over 800 substations data are collected by Western Power Distribution in South Wales, the UK, covering a good mix of industrial, commercial and domestic customers. For DER data, a set of 35 PVs panels are monitored over a year in the UK. Solar irradiance data are publicly available from Met Office based on the information of the region of the PVs. Real EVs charging data from the Low Carbon London (LCL) project [35] is used to develop typical EV charging profiles while EV charging data from the Winnipeg, Manitoba EV trials are used to test the proposed method.

Fig. 4 shows an example of aggregated regional load metered at a substation including 100 traditional customers, 10 PVs and 20 EVs over 10 days. The sampling time for all data are unified to half an hour, and 10-day data will be used as the basis for this study.

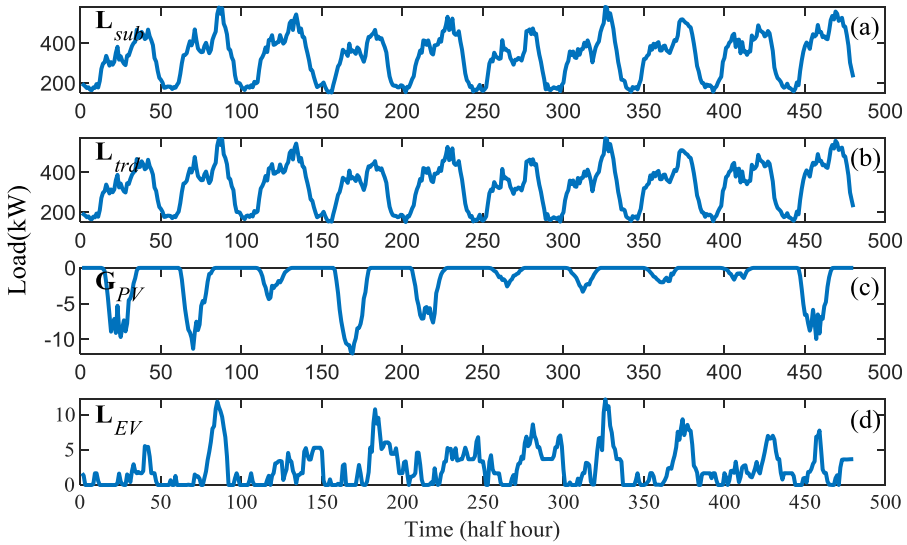

Fig. 4 Substation aggregated data (a) composited with traditional customers (b), PVs (c) and EVs (d).

3.1 Empirical mode decomposition denoise

As plotted in Fig. 5, *Residual* **A** *(red)* and $\mathbf{Err}_{est}$ *(blue)* overlap significantly in the time domain. However, in the frequency domain, the spectrum of $\mathbf{Err}_{est}$ *(blue)* is evenly distributed over the frequencies while the spectrum of *Residual* **A** *(red)* is skewed towards the low-frequency end.

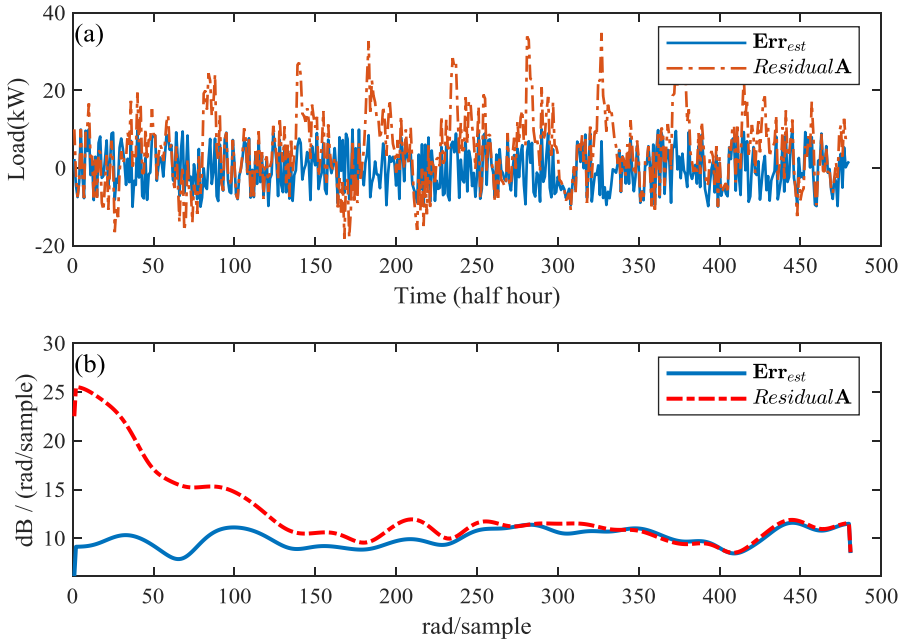

Fig. 5 Domain analysis of forecasting residual load for original and original added PVs & EVs load (a) time domain (b) frequency domain.

Fig. 6 shows the EMD decomposition results for *Residual* **A**. It can be seen that the high frequency part is mainly concentrated in the first IMF. Therefore, we can select u = 1 to reconstruct the *Residual* **A**.

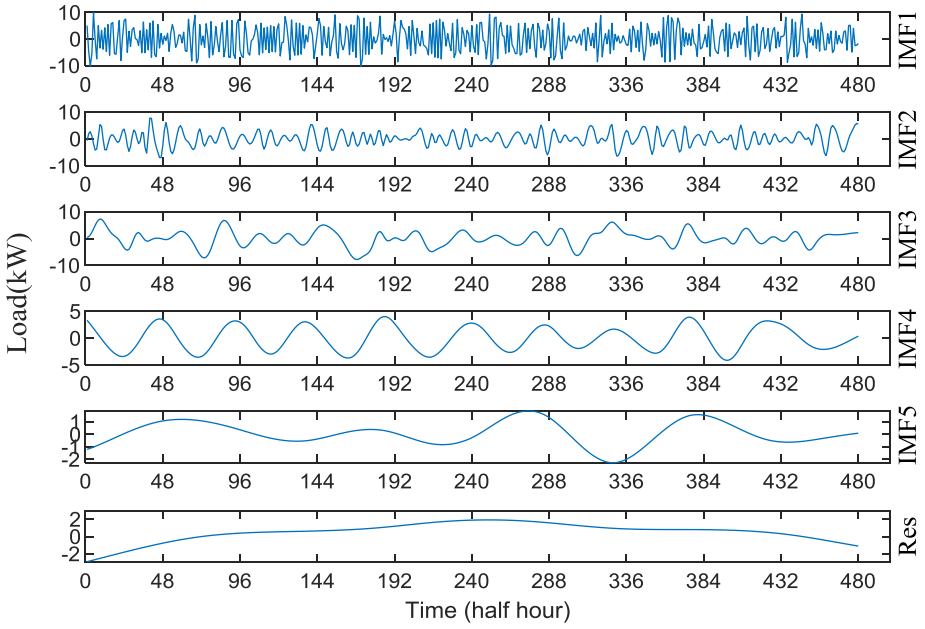

Fig. 6 EMD decompose of the *Residual A* into 5 IMFs and a residual.

In Fig. 7, the dashed (black) line is the raw signal of *Residual* **A**, and the solid (blue) line is the real summation load of PVs generation and the EVs load. There are clear deviations between the black and blue line due to the forecasting error. The EMD denoised signal, *Residual* $\mathbf{A}^*$, dotted (red) line is much closer to the expected signal (blue), matching well with the overall pattern and most peaks.

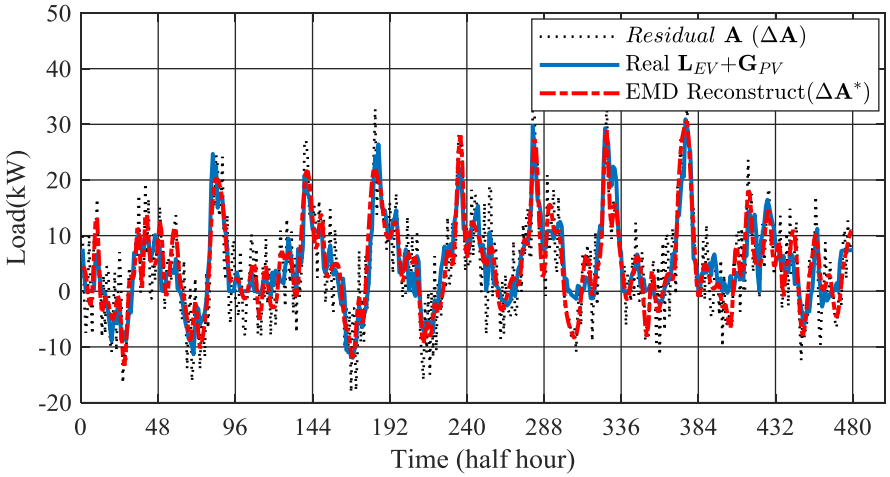

Fig. 7 The comparison of residual **A**, real EVs load added PVs generation and the EMD reconstruct signal.

3.2 Peak coincidence analysis for PV estimation

The solar energy output of a photovoltaic system is highly dependent on the solar irradiance. The upper graph in Fig. 8 shows the normalized solar irradiance and the bottom graph in Fig. 8 shows the energy output of 10 solar panels in the same region. It can be seen that the output of the solar panels is highly correlated with the intensity of solar radiation.

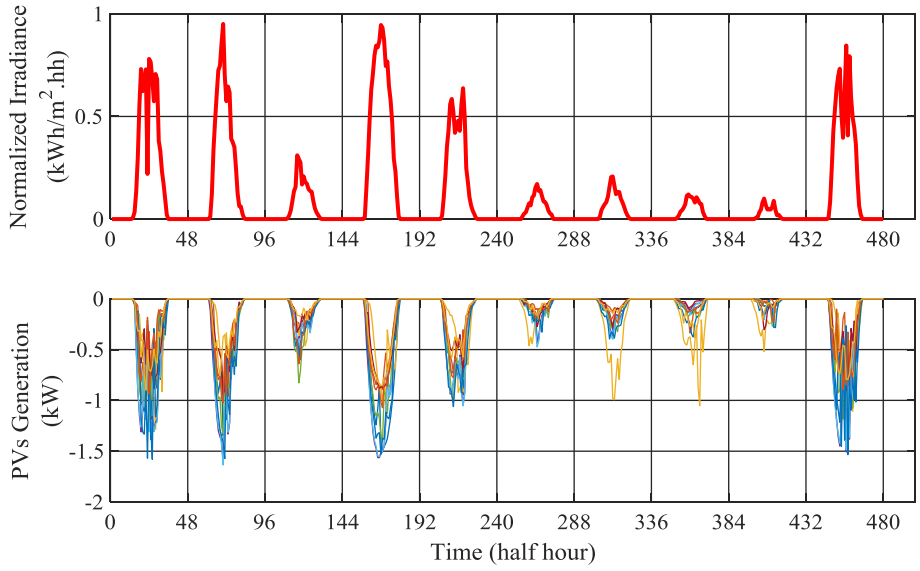

Fig. 8 Solar irradiance and 10 PV panels generation around Swansea area from Oct 3 2012 to Oct 12 2012.

3.3 Limited activation matching pursuit for electric vehicles estimation

In this study, real EV charging data from the Low Carbon London (LCL) project [35] are used. The load profiles cover 24 hours and the charging load is given a half hour time base as shown in Fig. 9 (one EV over one week). In order to make the algorithm converge as soon as possible, the correlation between atoms should be kept as low as possible. For example, orthogonal bases are used in orthogonal matching pursuit (OMP). However, in this study, it is difficult to generate an orthogonal basis without distortion on the of the EV's real charging patterns. Therefore, historical EV charging data are kept in their original form. Fig. 10 shows the selected atoms from the historical charging profiles of 60 EVs over 2 years. In this dictionary, we have 396 atoms, the correlation threshold τ is 0.85 and the minimum valid data is $e > 0.05$kW.

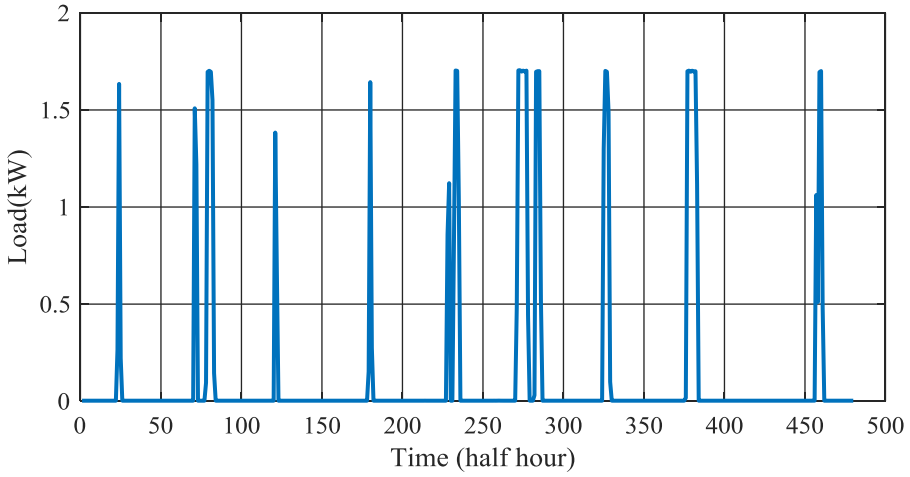

Fig. 9. The charging load profiles of EV from 13/10/2012 to 19/10/2012.

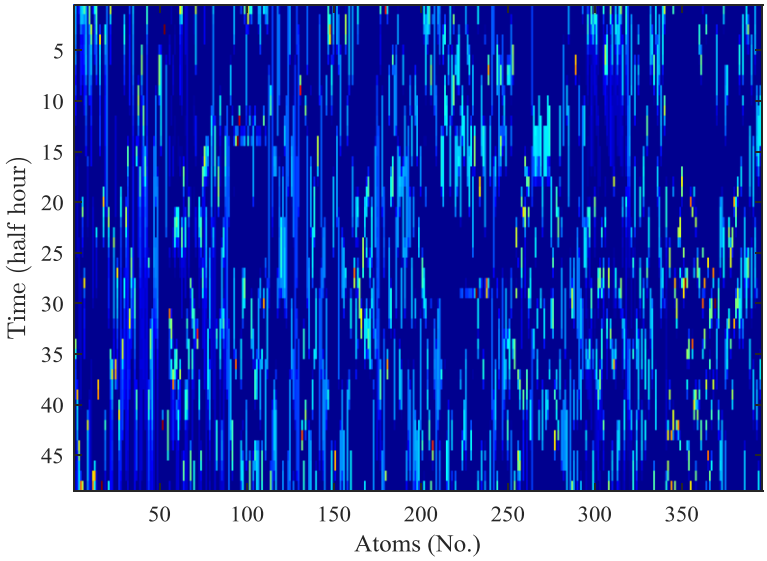

Fig. 10. A selected 396 atoms dictionary in day length and weak correlation (correlation threshold τ < 0.85, e > 0.05 kW).

There are two parameters need to be determined for the proposed LAMP method, the limited active coefficient c, and the termination condition. As the charging power of most EVs is below 10 kW [36], the maximum activation coefficient is set to 5, reflecting the maximum of 5 kWh of energy charged within half an hour. The iterations will be terminated when the root mean square error (RSME) improvement is less than a threshold, which is set to 0.005 in this study. As shown in Fig. 11, after 70 iterations, the RSME of the studied substation has converged at 0.92 kW.

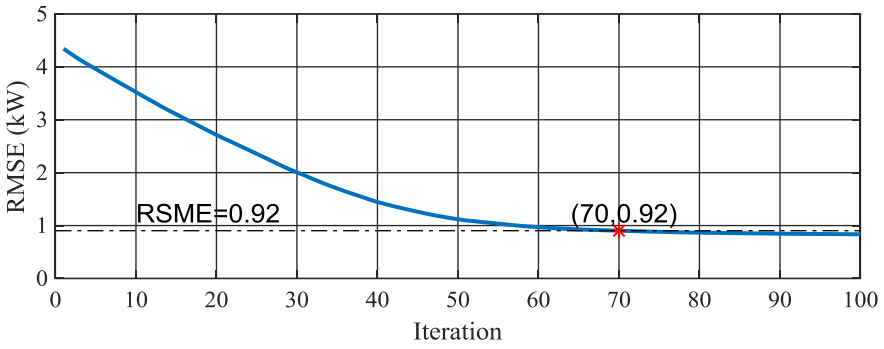

Fig. 11 Number of iteration and RMSE in a LAMP processing.

4. Results

4.1 Photovoltaic generate power estimating

The PVs generation energy can be estimated using equation (19). The *negative part* of the estimated DERs load ($\Delta\mathbf{A}^* < 0$) can be considered as an approximate estimate of the PVs' output power $\hat{\mathbf{G}}_{PV} \approx \Delta\mathbf{A}^* (\Delta\mathbf{A}^* < 0)$. The first 10 peaks of the output power $\hat{\mathbf{G}}_{PV}$ and solar irradiance are shown in Fig. 12. There are two points $\hat{G}$ (164, -10.93), $\hat{G}$ (169, -12.73) and H(165, 0.80), H(168, 0.93) that are matched to each other, and are marked with the symbol 'o'. According to (8), the coefficient $\hat{K}$ is -13.67 in this experiment.

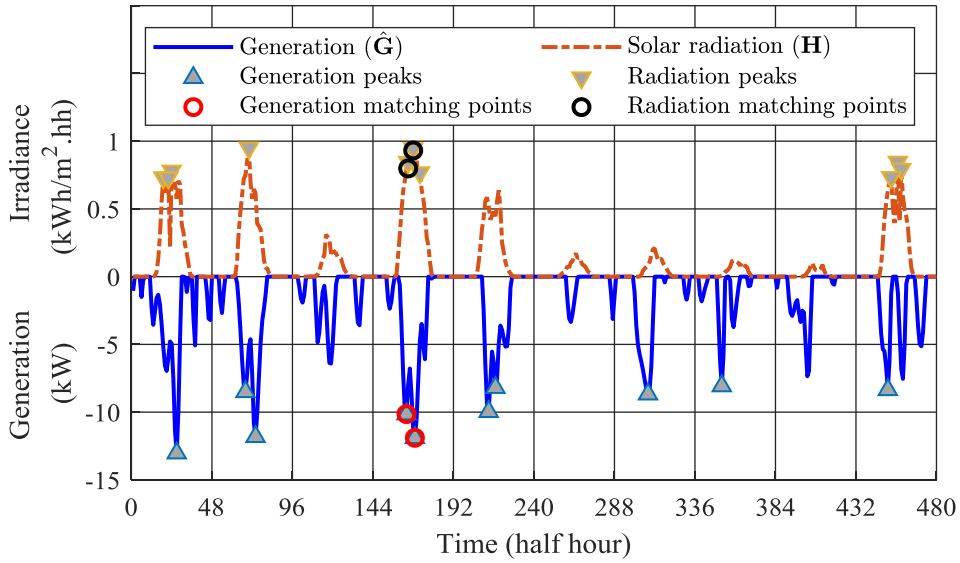

Fig. 12 Estimated solar energy generation coefficient K via matching the local maximum value.

The output energy of PVs is estimated with equation (10). In Fig. 13, the dotted (red) line is the estimated PVs' output energy, the solid (blue) line is the real PVs' output energy. The RSME of the estimated and real value of the PVs' output is 6.25 kW. The mean absolute percentage error (MAPE) of real PV and estimated is 19.40%, where the zero points of real PV is removed in the assessment of MAPE.

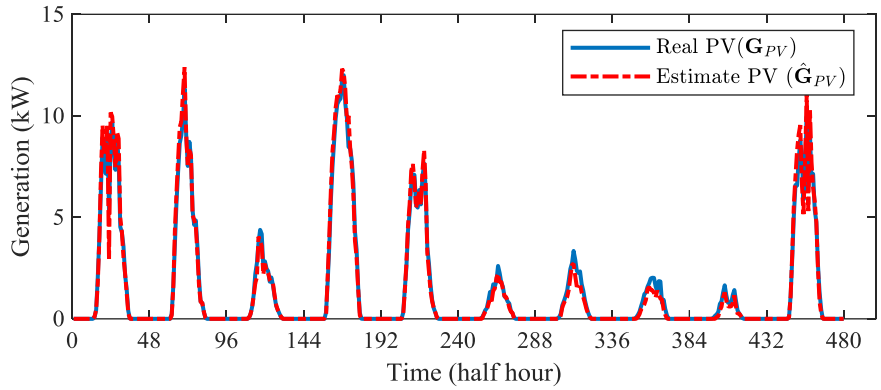

Fig. 13 Compared with real and estimate PVs output energy.

4.2 Electric vehicles load and number estimating

The estimated load of EVs $\hat{\mathbf{L}}_{EV}$ is the difference between $\Delta\mathbf{A}^{*}$ and $\hat{\mathbf{G}}_{PV}$, as described in (11). The estimated and real load of EVs are shown in Fig. 14 with dotted (red) and solid (blue) line, respectively. In Fig. 14, the estimated EV load (red) conforms well to the real EV load (blue) over the 10 days. There are a few noticeable overestimates at the start and end of the period. This is due to the end effect of EMD where the noise was not fully filtered at the two ends. The error is inherited from the previous processes of EMD denoise and PV estimation. The RSME between the estimated EVs load and the real EVs load is 4.01 kW.

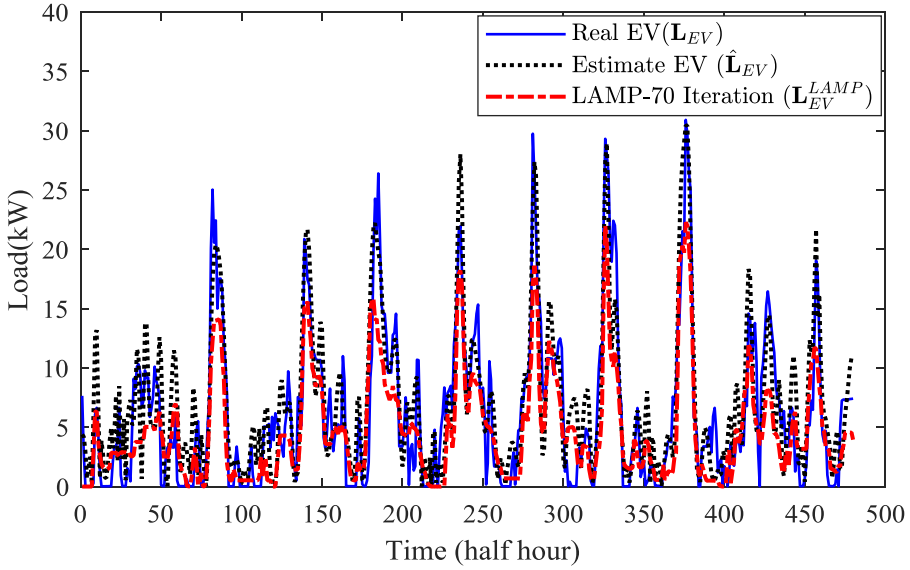

Fig. 14 Comparison between real (blue), estimated (black) and reconstructed (red) EV load.

The total EV load is then disaggregated into multiple EVs with different charging modes by the proposed LAMP method. To validate the atoms and coefficients derived, they are used to reconstruct the total EV load $\mathbf{L}_{EV}^{LAMP}$, which is depicted in the red dotted line in Fig. 14. The reconstructed EV profile follows the original EV profile very well, indicating a small residual. The RSME between $\mathbf{L}_{EV}^{LAMP}$ and $\hat{\mathbf{L}}_{EV}$ is 0.92 kW.

Since the overall load conforms, it is of interest to investigate whether the number of EVs can be accurately detected. The atom derived from LAMP indicates the charging mode and the activation coefficient indicates the number of EVs. The number of EVs charged at the substation is shown in Fig. 15, at every half an hour over 10 days. The red bars are the real

number of EVs while the blue bars are the estimation by LAMP. The estimation is relatively accurate with a mean absolute difference (MAD) of 1.17 EVs and a relative mean absolute difference (RMAD) of 55.77%. Due to the window effect in the sliding match, the estimation is very accurate for the middle 8 days, but higher deviation can be seen for the first and last day at both ends.

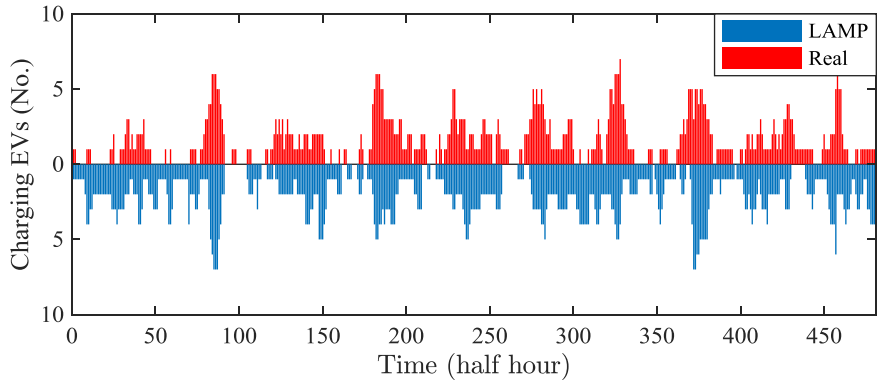

Fig. 15 Compare with real and LAMP estimated number of in charging EVs.

4.3 Comparison with classic nonintrusive load monitoring techniques

This section aims to compare the proposed LAMP with three classic NILM techniques: sparse coding (SC) [37] orthogonal matching pursuit (OMP [38]), and non-negative matching pursuit (NMP [39]). For details on the three classic techniques, please refer to Appendix A, B and C, respectively. In order to perform a fair comparison, the following modifications to improve the performance of classic techniques have been made: i) the same forecasting model is applied when estimating the traditional loads and EMD is implemented to reduce noise; ii) the same coefficients are used when estimating PV's generation; iii) all techniques use the same dictionary learned from the real EV data instead of unexplainable waveforms like sinusoids.

The reconstructions of the total EV loads are evaluated using MAPE. For the EV number estimation, four techniques are evaluated from two perspectives: i) reconstruction of total EV load measured by RSME, which is used for NILM [40]; and ii) the estimation of number of EV measured by the double-exponential (DE) score expressed in (32), which is popular evaluation index for forecasting competitions, using

$$score = \sum_{i=1}^{n} \exp^{-ln(1/2)\cdot|\Delta N_i/4|}, \tag{32}$$

where i is the time index and ΔN_i is the difference value between actually and estimated number of EVs at time i. Higher scores are awarded for small estimation errors to encourage accurate forecasting rather than mean-value forecasting. The comparison is performed on 27 substations with different forecasting errors, PV capacity and number of EVs. The test data configuration is listed in Table 1. The forecasting error is approximated at 1%, 3% and 5%. The number of PVs are 5, 10 and 20, and the number of EVs are 10, 20 and 30. There are a total of 27 cases in the comparison and the scores are listed for each case, with the best performing technique underlined. Overall, the LAMP method outperforms the other three techniques in the majority of cases. The advantage of using the LAMP becomes larger when the forecasting error and number of DERs increase. The LAMP performs consistently well with different forecasting errors and DERs penetration rates. The OMP only performs well when the forecasting error and DERs penetration rate are low. The likely reason is the

orthogonal constraints of OMP prevent it from local over-fitting only when the noise interference is low. The SC produces larger errors than LAMP and OMP, but lower errors than NMP.

Table 1. Comparative assessment between LAMP and classic NILM methods across 27 substations.

Case (No.)	Error (%)	PVs (No.)	EVs (No.)	RSME				Double Exponential (DE) Score			
				LAMP	SC	OMP	NMP	LAMP	SC	OMP	NMP
1	1	5	10	3.24	3.13	3.12	3.52	397.19	377.60	359.50	371.94
2	1	5	20	3.16	3.17	3.16	3.72	410.05	388.40	383.10	384.58
3	1	5	30	3.24	3.32	3.29	3.83	415.03	393.20	377.28	365.94
4	1	10	10	3.14	3.01	2.98	3.38	400.95	369.56	371.84	370.76
5	1	10	20	2.98	3.00	2.98	3.51	413.25	382.75	386.86	383.89
6	1	10	30	2.89	3.05	3.03	3.40	414.22	388.62	374.21	364.26
7	1	20	10	2.89	2.80	2.78	3.19	411.13	377.88	371.21	385.83
8	1	20	20	2.83	2.80	2.79	3.10	412.77	385.03	386.53	386.54
9	1	20	30	2.86	3.10	3.06	3.44	425.92	392.78	382.19	374.49
10	3	5	10	6.35	10.24	10.17	11.65	400.67	328.97	359.04	357.58
11	3	5	20	5.68	9.55	9.47	10.88	401.67	339.55	361.02	361.68
12	3	5	30	5.40	9.55	9.50	10.88	403.78	351.57	365.06	369.06
13	3	10	10	6.03	10.13	10.07	11.52	398.27	327.81	372.15	362.14
14	3	10	20	5.50	9.34	9.27	10.58	403.88	339.59	366.78	371.08
15	3	10	30	5.16	9.19	9.10	10.45	404.46	346.82	369.96	376.49
16	3	20	10	5.54	9.07	9.00	10.37	393.66	331.13	365.99	370.17
17	3	20	20	5.05	8.33	8.27	9.46	406.05	336.64	377.86	381.53
18	3	20	30	4.85	8.78	8.69	9.94	402.98	344.97	365.89	369.05
19	5	5	10	6.99	17.55	17.43	19.74	413.31	321.37	355.63	352.73
20	5	5	20	6.84	16.99	16.87	19.23	412.53	326.05	360.01	359.42
21	5	5	30	6.29	16.49	16.38	18.44	409.26	340.40	360.24	368.21
22	5	10	10	6.67	17.41	17.30	20.03	411.44	315.60	365.50	359.92
23	5	10	20	6.37	16.83	16.71	19.32	413.48	339.82	372.87	367.88
24	5	10	30	6.18	16.70	16.61	19.01	409.12	353.07	368.75	374.00
25	5	20	10	6.23	16.45	16.35	18.76	409.12	314.23	362.39	368.66
26	5	20	20	5.94	16.23	16.10	18.65	412.40	334.69	379.14	385.00
27	5	20	30	5.82	15.45	15.31	17.66	405.52	345.49	374.67	373.87

In terms of the DE score, the LAMP consistently performs best while sparse coding has the lowest DE score. The DE scores of NMP and OMP are relatively close, between 360 and 400, decreasing when the forecasting error increases. When the error increases to 3% and 5%, the DE scores of SC decrease rapidly, indicating it is sensitive to noise. The performances of NMP,

OMP and SC all degrade when the number of EVs increases while the impact of PV is less significant. Among the four algorithms, LAMP all shows the highest score through all 27 cases and is robust when there are additional forecasting noise, EVs and PVs. In order to verify the stability of the LAMP under the scenario of high EV penetration, a case study of 50 EVs is conducted. The result in Fig. 16 shows that when the forecasting error level is 1%, the DE score of LAMP increases within the number of EVs. As the forecasting error increasing to 3% and 5%, the LAMP's DE scores decrease as the number of EVs increases. However, the DE score of LAMP is always stable at around 390, which is 30 points higher than other algorithms.

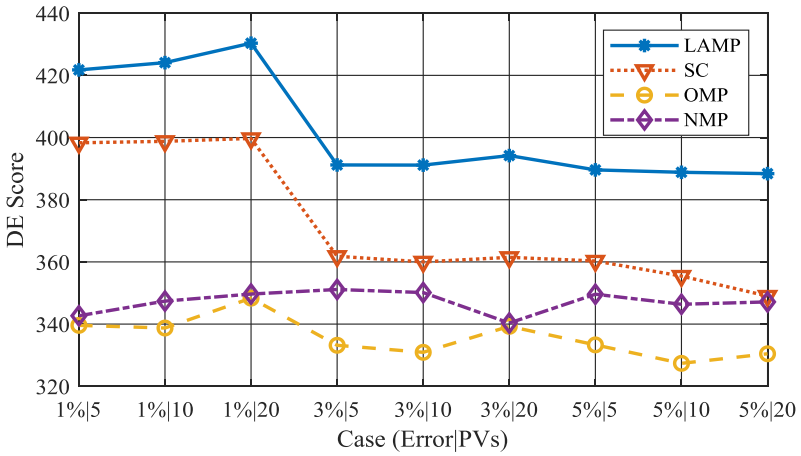

Fig. 16 The scores of LAMP, SC, OMP and NMP algorithms under different forecasting errors and PV numbers with 50 EVs.

Fig. 17 shows an example of the number of EVs estimated by different algorithms for Case 14 in Table 1. These results show that LAMP can capture the overall trend of the EV number variation and is the best performing method for estimating the local maximum and minimum values. Estimations of sparse coding are much higher than real values. This is because of the sparse limitation restricts the number of non-zero values in the sparse activation coefficients and thus makes each value higher than the real case. The estimations of OMP and NMP are flat over time. This might be caused by the greedy algorithms try to fit the load curve as much as possible without considering the real world constraints and thus unable to catch the spikes.

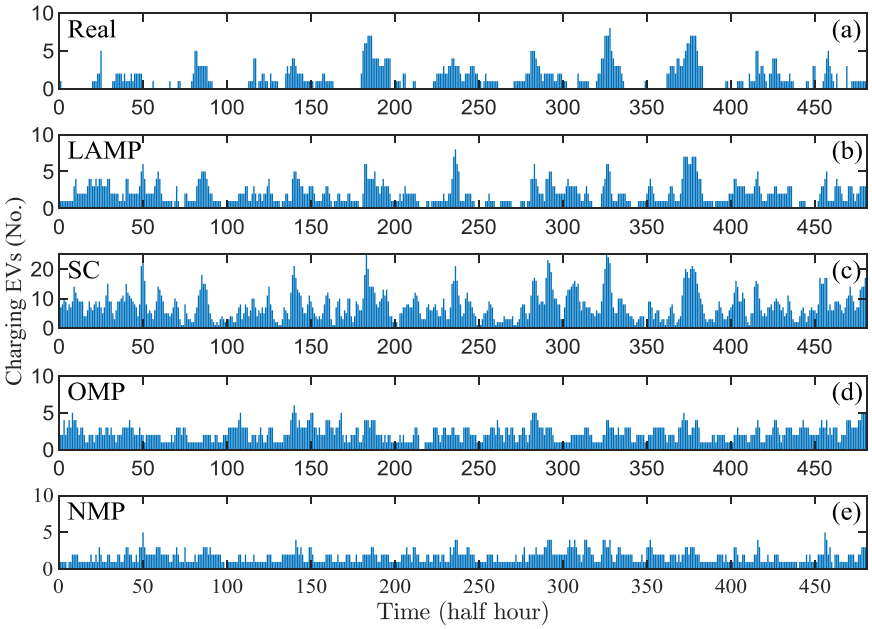

Fig. 17 Real Number of charging EVs (a) and estimation by LAMP (b), SC (c), OMP (d) and NMP (d) for case 14(error=3%, PVs=10, EVs=20).

5. Conclusion

In this paper, an upgraded NILM problem is proposed that is able to disaggregate regional loads (e.g. traditional load, PVs and EVs) rather than just household appliances. This is a critical research topic to address the challenges the increasing scale of DERs present to the traditional operation of distribution networks. The proposed Regional-NILM method provides a way to estimate the real-time status of DERs without expensive monitoring and data privacy issues.

The proposed three-stage approach firstly estimates the traditional load using forecasting models. The unbiased forecasting errors are filtered using EMD reconstruction without the high-frequency IMFs. Peak coincidence analysis is then performed between negative residuals and solar irradiance to identify the installed PV capacity and related coefficients. Finally, a novel LAMP method is proposed to analyse the final residual for the number of EVs and total charging load.

The method is validated using 27 cases based on real data from traditional substations, EV and PV trials in the UK. In particular, 27 substations are tested with different penetration rates of DERs and forecasting errors. Results show the regional aggregated load can be accurately decomposed to the traditional load, PV generation and EV load. For EV number estimation, the proposed LAMP achieved an improved performance over that of classic SC, OMP and NMP techniques by 16.53%, 10.21% and 10.00%, respectively. The improvement becomes more significant (to 23.79%, 12.08% and 11.75%, respectively) when the substation has a higher DERs' penetration rate and a larger forecasting error.

This paper focuses on developing the overall methodology of regional-NILM. Future research could investigate different enabling techniques such as wavelet analysis for noise reduction, a deep neural network for EV and PV profiles extraction and hidden Markov model (HMM) for EV charging mode estimation.

6. References

[1] National-Grid. Future energy scenarios (2017), http://fes.nationalgrid.com/fes-document/; 2017 [accessed 15 Jan 2019].

[2] Asrari A, Wu T, Lotfifard S. The Impacts of Distributed Energy Sources on Distribution Network Reconfiguration. IEEE Transactions on Energy Conversion. 2016;31:606-13.https://doi.org/10.1109/TEC.2015.2514191.

[3] Chowdhury S, Crossley P. Microgrids and active distribution networks: The Institution of Engineering and Technology; 2009.

[4] Nosratabadi SM, Hooshmand R-A, Gholipour E. A comprehensive review on microgrid and virtual power plant concepts employed for distributed energy resources scheduling in power systems. Renewable and Sustainable Energy Reviews. 2017;67:341-63.https://doi.org/10.1016/j.rser.2016.09.025.

[5] Wu Z, Tazvinga H, Xia X. Demand side management of photovoltaic-battery hybrid system. Applied Energy. 2015;148:294-304.https://doi.org/10.1016/j.apenergy.2015.03.109.

[6] Mengelkamp E, Gärttner J, Rock K, Kessler S, Orsini L, Weinhardt C. Designing microgrid energy markets: A case study: The Brooklyn Microgrid. Applied Energy. 2018;210:870-80.https://doi.org/10.1016/j.apenergy.2017.06.054.

[7] Qi W, Zhang C, Yi D, Xydis G, Wang J, Østergaard J. Review of real-time electricity markets for integrating Distributed Energy Resources and Demand Response. Applied Energy. 2015;138:695-706.https://doi.org/10.1016/j.egypro.2019.01.040.

[8] Gonçalves H, Ocneanu A, Bergés M. Unsupervised disaggregation of appliances using aggregated consumption data. The 1st KDD Workshop on Data Mining Applications in Sustainability (SustKDD) 2011.

[9] Hart GW. Nonintrusive appliance load monitoring. Proceedings of the IEEE. 1992;80:1870-91.https://doi.org/10.1109/5.192069.

[10] Esa NF, Abdullah MP, Hassan MY. A review disaggregation method in Non-intrusive Appliance Load Monitoring. Renewable and Sustainable Energy Reviews. 2016;66:163-73.https://doi.org/10.1016/j.rser.2016.07.009.

[11] Cominola A, Giuliani M, Piga D, Castelletti A, Rizzoli AE. A hybrid signature-based iterative disaggregation algorithm for non-intrusive load monitoring. Applied energy. 2017;185:331-44.https://doi.org/10.1016/j.apenergy.2016.10.040.

[12] Welikala S, Dinesh C, Ekanayake MPB, Godaliyadda RI, Ekanayake J. Incorporating Appliance Usage Patterns for Non-Intrusive Load Monitoring and Load Forecasting. IEEE Transactions on Smart Grid. 2017.https://doi.org/10.1109/TSG.2017.2743760.

[13] Norford LK, Leeb SB. Non-intrusive electrical load monitoring in commercial buildings based on steady-state and transient load-detection algorithms. Energy and Buildings. 1996;24:51-64. https://doi.org/10.1016/0378-7788(95)00958-2.

[14] Enríquez R, Jiménez M, Heras M. Towards non-intrusive thermal load Monitoring of buildings: BES calibration. Applied energy. 2017;191:44-54.https://doi.org/10.1016/j.apenergy.2017.01.050.

[15] Yue X, Liu J, Ran L, Li F, Gu C, Tang S. Economic planning of electric vehicle charging stations considering traffic constraints and load profile templates. Applied Energy. 2016;178:647-59. https://doi.org/10.1016/j.apenergy.2016.06.021.

[16] Dinesh C, Welikala S, Liyanage Y, Ekanayake MPB, Godaliyadda RI, Ekanayake J. Non-intrusive load monitoring under residential solar power influx. Applied Energy. 2017;205:1068-80.https://doi.org/10.1016/j.apenergy.2017.08.094.

[17] Bo L, Luan W, Yu Y. Dynamic time warping based non-intrusive load transient identification. Applied Energy. 2017;195:634-45.https://doi.org/10.1016/j.apenergy.2017.03.010.

[18] Aboulian A, Green DH, Switzer JF, Kane TJ, Bredariol GV, Lindahl P, et al. NILM Dashboard: A Power System Monitor for Electromechanical Equipment Diagnostics. IEEE Transactions on Industrial Informatics. 2018;PP:1-.https://doi.org/10.1109/TII.2018.2843770.

[19] Stankovic L, Stankovic V, Liao J, Wilson C. Measuring the energy intensity of domestic activities from smart meter data. Applied Energy. 2016;183:1565-80.https://doi.org/10.1016/j.apenergy.2016.09.087.

[20] Welikala S, Thelasingha N, Akram M, Ekanayake PB, Godaliyadda RI, Ekanayake JB. Implementation of a robust real-time non-intrusive load monitoring solution. Applied Energy. 2019;238:1519-29.https://doi.org/10.1016/j.apenergy.2019.01.167.

[21] Rashid H, Singh P, Stankovic V, Stankovic L. Can non-intrusive load monitoring be used for identifying an appliance's anomalous behaviour? Applied Energy. 2019;238:796-805.https://doi.org/10.1016/j.apenergy.2019.01.061.

[22] Shuangyuan W, Ran L, Adrian E, Furong L. Electric Vehicle Load Disaggregation Based on Limited Activation Matching Pursuits. Proc 10th International Conference on Applied Energy (ICAE). Hong Kong, China2018.

[23] ELEXON. Load Profiles and their use in Electricity Settlement, https://www.elexon.co.uk/wp-content/uploads/2013/11/load_profiles_v2.0_cgi.pdf; 2013[accessed 01 March 2019].

[24] Goude Y, Nedellec R, Kong N. Local short and middle term electricity load forecasting with semi-parametric additive models. IEEE transactions on smart grid. 2014;5:440-6.https://doi.org/10.1109/TSG.2013.2278425.

[25] Livera AMD, Hyndman RJ, Snyder RD. Forecasting Time Series With Complex Seasonal Patterns Using Exponential Smoothing. Monash Econometrics & Business Statistics Working Papers. 2011;106:1513-27.https://doi.org/10.1198/jasa.2011.tm09771.

[26] Shi H, Xu M, Ran L. Deep Learning for Household Load Forecasting – A Novel Pooling Deep RNN. IEEE Transactions on Smart Grid. 2017;PP:5271 - 80.https://doi.org/10.1109/TSG.2017.2686012.

[27] Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society of London Series A: Mathematical, Physical and Engineering Sciences. 1998;454:903-95

[28] Boudraa AO, Cexus JC. EMD-Based signal filtering. IEEE Transactions on Instrumentation & Measurement. 2007;56:2196-202.https://doi.org/10.1109/TIM.2007.907967.

[29] Junsheng C, Dejie Y, Yu Y. A fault diagnosis approach for roller bearings based on EMD method and AR model. Mechanical Systems and Signal Processing. 2006;20:350-62.https://doi.org/10.1016/j.ymssp.2004.11.002.

[30] Photovoltaic-software. How to calculate output energy of PV solar systems?, https://photovoltaic-software.com/principle-ressources/how-calculate-solar-energy-power-pv-systems; 2013 [accessed 01 March 2019].

[31] Wang S, Du L, Ye J, Zhao D. Robust identification of ev charging profiles. 2018 IEEE Transportation Electrification Conference and Expo (ITEC): IEEE; 2018. p. 1-6.

[32] Mallat SG, Zhang Z. Matching pursuits with time-frequency dictionaries. IEEE Transactions on signal processing. 1993;41:3397-415.https://doi.org/10.1109/78.258082.

[33] Hartigan JA, Wong MA. Algorithm AS 136: A K-Means Clustering Algorithm. Journal of the Royal Statistical Society. 1979;28:100-8

[34] Grosse R, Raina R, Kwong H, Ng AY. Shift-invariance sparse coding for audio classification. arXiv preprint arXiv:12065241. 2012

[35] UK-Power-Networks. Low Carbon London Project, http://innovation.ukpowernetworks.co.uk/innovation/en/Projects/tier-2-projects/Low-Carbon-London-(LCL); 2012 [accessed 01 March 2019].

[36] Munshi AA, Mohamed YA-RI. Extracting and Defining Flexibility of Residential Electrical Vehicle Charging Loads. IEEE Transactions on Industrial Informatics. 2018;14:448-61.https://doi.org/10.1109/TII.2017.2724559.

[37] Liu J, Ji S, Ye J. SLEP: Sparse learning with efficient projections. Arizona State University. 2009;6:7

[38] Cai TT, Wang L. Orthogonal Matching Pursuit for Sparse Signal Recovery With Noise. IEEE Transactions on Information Theory. 2011;7:4680-8.https://doi.org/10.1109/TIT.2011.2146090.

[39] Yaghoobi M, Wu D, Davies ME. Fast Non-Negative Orthogonal Matching Pursuit. IEEE Signal Processing Letters. 2015;22:1229-33.https://10.1109/LSP.2015.2393637.

[40] Zoha A, Gluhak A, Imran MA, Rajasegarar S. Non-intrusive load monitoring approaches for disaggregated energy sensing: A survey. Sensors. 2012;12:16838-66.https://doi.org/10.3390/s121216838.

7. Appendix

A. Sparse Coding (SC)

The sparse coding approach usually including two stages: training and coding. In the training stage a dictionary will be learnt from the training samples. However, the atoms in the dictionary is consist of the energy usage of EVs in this paper. Instead of learning, we constructing a dictionary that preserves the charging characteristics of electric vehicles. Therefore, here we mainly use the coding stage of the SC. In the coding stage, we can compute the representation coefficients one by one independently through the following minimization problem.

The classic approach is to solve the least absolute shrinkage and selection operator (LASSO) problem:

$$\min_{\mathbf{W}} \|\mathbf{X}-\mathbf{B}\mathbf{W}\|^2 + \beta|\mathbf{W}|, \tag{A.1}$$

where, $\mathbf{W}$ is the representation coefficients, $\mathbf{B}$ is the constructed dictionary, $\mathbf{X}$ is the test signal, and β is the sparse parameter $|\mathbf{W}|$ denote the L1-norm. This paper implements SC based on the references [37].

B. Orthogonal Matching Pursuit (OMP)

The Orthogonal Matching Pursuit Algorithm (OMP) [38] is a greedy compressed signal recovery algorithm which selects the best fitting atoms of the dictionary in each iteration. Matching Pursuit (MP) can only guarantee that the residual vector is orthogonal to the selected atoms of the dictionary in each iteration, but is generally not orthogonal to the previously selected atoms. OMP can ensure that the residual vector is orthogonal to all the previously selected atoms after each iteration to ensure the optimality of the iteration, thereby reducing

the number of iterations and the performance is more robust. The calculation steps of OMP can be summarized as follows:

Input: observed signal vector $\mathbf{X} \in R^m$ and dictionary $\mathbf{B} \in R^{m \times n}$

Output: sparse coefficients vector $\mathbf{W} \in R^n$

Step 1. Initialization, let the label set $T_0 = \varnothing$, the initial residual vector $r_0 = \mathbf{X}$, $k = 1$.

Step 2. Identification, find the atom in dictionary **B** that has the strongest correlation with the residual vector r_{k-1}, $j_k \in \arg\max \mathbf{W}_j \left| < r_{k-1}, \mathbf{B}_j > \right|$, $T_k = T_{k-1} \bigcup$.

Step 3. Estimation, minimize the problem $\arg min_{\mathbf{W}} = \| \mathbf{X} - \mathbf{B}_{T_k} \mathbf{W} \|$ by $\mathbf{W}_k = (\mathbf{B}_{T_k}^H \mathbf{B}_{T_k})^{-1} \mathbf{B}_{T_k}^H \mathbf{X}$, where $\mathbf{B}_{T_k} = [\mathbf{b}_{t_1}, \cdots$, $t_1, \cdots$, $\mathbf{B}^H$ is conjugate transpose

Step 4. Update the residual $r_k = \mathbf{X} - \mathbf{B}_{T_k} \mathbf{W}_k$.

Step 5. Let $k = k + 1$, and repeat Step 2 to Step 4, if a stop criterion is met, the iteration is stopped.

Step 6. Output the coefficients vector $\mathbf{W}(i) = \mathbf{W}_k(i), if\ i \in T_k$; $\mathbf{W}(i) = 0, otherwise$.

Here are three common stop criteria:

1) Stop after running to a fixed number of iterations.

2) The residual energy is less than a predetermined value δ, $\|r_k\|_2 \le \delta$

3) When any atom of the dictionary **B** has no significant energy of the residual vector r_k, $\left\|\mathbf{B}^H r_k\right\|_\infty \le \delta$

C. Non-negative Matching Pursuit (NMP)

In order to obtain non-negative coefficients, we need some modifications in MP or OMP. At each iteration only the most "positively" correlated atom will be selected, then only the positive coefficients could be obtained [39]. A sub-optimal approach named Non-Negative Least Square (NNLS) program is used to refine the selected coefficients at each iteration in OMP Step 3, which will change to $\arg\min_{\mathbf{W} \ge 0} = \| \mathbf{X} - \mathbf{B}_{T_k} \mathbf{W} \|$.